```
OM protein - protein search, using sw model

                        GenCore version 5.1.7
           Copyright (c) 1993 - 2006  Bioacceleration Ltd.

Run on:            March 13, 2006, 18:37:13 ; Search time 167 Seconds
                                               (without alignments)
                                               950.749 Million cell updates/sec

Title:             US-09-931-701A-2
Perfect score:     1901
Sequence:          1 MNKPLGKIVASTALLISVAF.............SLGSTNLYGSGLVNABAATR 380

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters:        1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 listing first 45 summaries

Database :         Published Applications AA Main:*
                   1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                   2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                   3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                   4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                   5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                   6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                SUMMARIES

Result          %       Query
No.     Score   Match   Length   DB   DB ID                     Description
------  -----   -----   ------   --   -----                     -----------
 1      1901    100.0   380      4    US-09-931-701-2           Sequence 2, Appli
 2      1777    93.5    380      4    US-10-324-152-8           Sequence 8, Appli
 3      1772    93.2    380      4    US-10-324-152-9           Sequence 9, Appli
 4      1770    93.1    380      4    US-09-824-893A-261        Sequence 261, Appl
 5      1770    93.1    380      4    US-10-736-997-261         Sequence 261, App
 6      1760    92.6    380      5    US-10-873-917-2           Sequence 2, Appli
 7      1754.5  92.3    377      5    US-09-813-408-1           Sequence 1, Appli
 8      1670.5  87.9    380      5    US-10-476-463-4           Sequence 4, Appli
 9      1667.5  87.7    380      5    US-10-836-959-4           Sequence 4, Appli
10      1665.5  87.6    380      5    US-10-836-959-2           Sequence 2, Appli
11      1665.5  87.6    380      5    US-10-476-463-2           Sequence 2, Appli
12      1664.5  87.6    380      5    US-10-836-959-6           Sequence 6, Appli
13      1490    78.4    374      5    US-10-873-610-2           Sequence 2, Appli
14      1461.5  76.9    379      4    US-09-813-408-6           Sequence 6, Appli
15      1387    73.0    382      3    US-09-813-408-7           Sequence 7, Appli
16      1311.5  69.0    383      5    US-10-872-162-2           Sequence 2, Appli
17      1303.5  68.6    383      5    US-10-872-166-2           Sequence 2, Appli
18      1284    67.5    269      3    US-09-837-235-16          Sequence 16, Appl
19      1284    67.5    269      4    US-10-336-324-4.          Sequence 4, Appli
20      1284    67.5    269      4    US-10-403-105-7           Sequence 7, Appli
21      1284    67.5    269      6    US-11-043-731-16          Sequence 16, Appl
22      1284    67.5    268      5    US-11-104-845-4           Sequence 4, Appli
23      1280    67.3    268      5    US-10-873-917-5           Sequence 5, Appli
24      1280    67.3    268      5    US-10-872-162-4           Sequence 4, Appli
25      1279    67.3    269      5    US-08-322-678-10          Sequence 10, Appl
26      1279    67.3    269      3    US-09-060-854B-6          Sequence 6, Appli
27      1279    67.3    269      3    US-09-975-139-1           Sequence 1, Appl
28      1279    67.3    269      3    US-09-976-414-8           Sequence 8, Appli
29      1279    67.3    269      3    US-09-736-116-49          Sequence 49, Appl
30      1279    67.3    269      3    US-09-957-806A-24         Sequence 24, Appl
31      1279    67.3    269      3    US-09-957-806A-56         Sequence 56, Appl
32      1279    67.3    269      4    US-10-075-907-1           Sequence 1, Appli
33      1279    67.3    269      4    US-10-075-895-1           Sequence 1, Appli
34      1279    67.3    269      4    US-10-033-325-6           Sequence 6, Appli
35      1279    67.3    269      4    US-10-209-812-3           Sequence 3, Appli
36      1279    67.3    269      4    US-10-104-693-4           Sequence 4, Appli
37      1279    67.3    269      4    US-10-228-572-6           Sequence 6, Appli
38      1279    67.3    269      4    US-10-242-549-5           Sequence 5, Appli
39      1279    67.3    269      4    US-10-324-152-5           Sequence 5, Appli
40      1279    67.3    269      4    US-10-313-853-1           Sequence 1, Appli
41      1279    67.3    269      4    US-10-336-324-7           Sequence 7, Appli
42      1279    67.3    269      4    US-10-403-105-10          Sequence 10, Appl
43      1279    67.3    269      4    US-10-243-576-5           Sequence 5, Appli
44      1279    67.3    269      5    US-10-423-649-6           Sequence 6, Appli
45      1279    67.3    269      4    US-10-323-324-10          Sequence 10, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-931-701-2
; Sequence 2, Application US/09931701
; Publication No. US20040241820A1
; GENERAL INFORMATION:
;  APPLICANT: Outtrup, Helle
;  APPLICANT: Pedersen, Poul
;  APPLICANT: Sorensen, Marianne
;  TITLE OF INVENTION: Subtilase Enzymes
;  FILE REFERENCE: 10065.200-US
;  CURRENT APPLICATION NUMBER: US/09/931,701
;  CURRENT FILING DATE: 2001-08-16
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
;
;  SEQ ID NO 2
;  LENGTH: 380
;  TYPE: PRT
;  ORGANISM: Bacillus clausii
US-09-931-701-2

  Query Match             100.0%;  Score 1901; DB 3; Length 380;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-138;
  Matches  380;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 MNKPLGKIVASTALLISVAFSSSIASAAEAKEKYLIGFNEQEAVSEFVEQVDANNDAV 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MNKPLGKIVASTALLISVAFSSSIASAAEAKEKYLIGFNEQEAVSEFVEQVDANNDAV 60

Qy         61 LSEEEEVEIELLHEFETIPVLSVELSPEDVDTLELDPAISYIEEDVEVSIMAQSVPWGIS 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LSEEEEVEIELLHEFETIPVLSVELSPEDVDTLELDPAISYIEEDVEVSIMAQSVPWGIS 120

Qy        121 RVQAPAAHNRGVTGSGVKVAVLDTGISAHPDLNIRGGASFVTGEPPTYQDGNGHGTHVAGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RVQAPAAHNRGVTGSGVKVAVLDTGISAHPDLNIRGGASFVTGEPPTYQDGNGHGTHVAGT 180

Qy        181 IAALNNSIGVLGVAPNAELYAVKVLAANGRGPVSSIAQGLEWAGNNGMDVANLSLGSPSP 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IAALNNSIGVLGVAPNAELYAVKVLAANGRGPVSSIAQGLEWAGNNGMDVANLSLGSPSP 240

Qy        241 SATLEQAVNSATSRGVLVVAATGNSGTGSLDYPARYANAMAVGATDQNNNRASFSQYGAG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SATLEQAVNSATSRGVLVVAATGNSGTGSLDYPARYANAMAVGATDQNNNRASFSQYGAG 300

Qy        301 LDIVAPGVNVQSTYPGSTYASFNGTSMATPHVVGVAALVKQKNPSWSNVQIRNHLKNTAT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LDIVAPGVNVQSTYPGSTYASFNGTSMATPHVVGVAALVKQKNPSWSNVQIRNHLKNTAT 360

Qy        361 SLGSTNLYGSGLVNABAATR 380
              ||||||||||||||||||||
Db        361 SLGSTNLYGSGLVNABAATR 380
```

OM protein - protein search, using sw model

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: March 13, 2006, 18:37:57 ; Search time 23 Seconds
(without alignments)
459.881 Million cell updates/sec

Title: US-09-931-701A-2
Perfect score: 1901
Sequence: 1 MNKPLGKIVASTALLISVAF........SLGSTNLYGSGLVNAEAATR 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Published Applications AA New:*
1: /cgn2_6/ptcdata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptcdata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptcdata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptcdata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptcdata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptcdata/1/pubpaa/US0_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1279	67.3	269	7	US-11-020-602-6	Sequence 6, Appli
2	972.5	51.2	272	7	US-11-020-602-236	Sequence 236, App
3	914	48.1	382	7	US-11-020-602-2	Sequence 2, Appli
4	898.5	47.3	379	7	US-11-156-062-23	Sequence 23, Appl
5	849	44.7	275	7	US-11-020-602-4	Sequence 4, Appli
6	835	43.9	275	7	US-11-065-943-54	Sequence 54, Appli
7	835	43.9	275	7	US-11-020-602-3	Sequence 3, Appli
8	815.5	42.9	274	7	US-11-156-062-14	Sequence 14, Appl
9	815.5	42.9	274	7	US-11-156-062-16	Sequence 16, Appl
10	814.5	42.8	274	7	US-11-020-602-5	Sequence 5, Appli
11	810.5	42.6	274	7	US-11-156-062-4	Sequence 4, Appli
12	809.5	42.6	274	7	US-11-156-062-10	Sequence 10, Appl
13	809.5	42.6	274	7	US-11-156-062-12	Sequence 12, Appl
14	809.5	42.6	274	7	US-11-156-062-18	Sequence 18, Appl
15	809.5	42.6	274	7	US-11-156-062-2	Sequence 2, Appli
16	804.5	42.3	274	7	US-11-156-062-6	Sequence 6, Appli
17	803.5	42.3	274	7	US-11-156-062-8	Sequence 8, Appli
18	609.5	32.1	591	6	US-10-510-386-22	Sequence 22, Appl
19	410	21.6	874	6	US-10-510-386-28	Sequence 28, Appl
20	410	21.6	1047	6	US-10-510-386-200	Sequence 200, Appl
21	355	18.7	802	6	US-11-020-602-209	Sequence 209, App
22	354	18.6	1432	6	US-10-510-386-218	Sequence 218, App
23	290	15.3	1052	7	US-11-020-602-208	Sequence 208, App
24	280.5	14.8	280	7	US-11-020-602-209	Sequence 209, App
25	234	12.3	49	7	US-11-011-666-12	Sequence 12, Appl
26	222	11.7		7	US-11-011-666-9	Sequence 9, Appli
27	222	11.7		7	US-11-096-568A-24714	Sequence 24714, A
28	222	11.7		7	US-11-096-568A-24713	Sequence 24713, A
29	222	11.7		7	US-11-096-568A-24712	Sequence 24712, A
30	222	11.7		7	US-10-501-035-361	Sequence 361, App
31	221	11.6		7	US-11-011-666-8	Sequence 8, Appli
32	211.5	11.1		7	US-11-096-568A-24027	Sequence 24027, A
33	211.5	11.1		7	US-11-096-568A-24026	Sequence 24026, A
34	210	11.0		7	US-11-096-568A-24026	Sequence 24026, A
35	210	11.0		6	US-10-218-986-2	Sequence 2, Appli
36	209	11.0		6	US-11-096-568A-24028	Sequence 24028, A
37	205.5	10.8		7	US-11-096-568A-31251	Sequence 31251, A
38	203	10.7		7	US-11-096-568A-31862	Sequence 31862, A
39	203	10.7		7	US-11-096-568A-31861	Sequence 31861, A
40	200.5	10.5		7	US-11-052-554A-260	Sequence 260, App
41	193.5	10.2		7	US-11-037-243-82	Sequence 82, Appl
42	192.5	10.1		7	US-11-096-568A-31863	Sequence 31863, A
43	190	10.0		7	US-11-096-568A-23983	Sequence 23983, A
44	190	10.0		7	US-11-096-568A-23982	Sequence 23982, A
45	187.5	9.9		6	US-10-873-528-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-11-020-602-6

; Sequence 6, Application US/11020602
; Publication No. US2006002476A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David
; APPLICANT: Harding, Fiona
; TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; FILE REFERENCE: GC527C2
; CURRENT APPLICATION NUMBER: US/11/020,602.
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: US 09/500,135
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: US 09/060,872
; PRIOR FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 6
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Bacillus lentus
US-11-020-602-6

Query Match   67.3%; Score 1279; DB 7; Length 269;
Best Local Similarity   93.3%; Pred. No. 5.7e-85;
Matches 251; Conservative   6; Mismatches 12; Indels   0; Gaps   0;

```
Qy    112 AQSVPWGISRVQAPAAHNRGVTGSGVKVAVLDTGISAHPDLNIRGGASFVTGEPTYQDGN 171
              ||||||||||||||||||| |||||||||||||||:||||||||||||| |||| ||
Db      1 AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGN 60

Qy    172 GHGTHVAGTIAALNNSIGVLGVAPNAELYAVKVLAAANGRGPVSSIAQGLEWAGNNGMDVA 231
              |||||||||||||||||||||||| ||||||||||| :||||||||||||||||||||
Db     61 GHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVA 120

Qy    232 NLSLGSPSPSATLEQAVNSATSRGVLVVAATGNSGTGSLDYPARYANAMAVGATDQNNNR 291
              |||||||||||||||||||||||||||:|||| :|| |  :|||||||||||||||||
Db    121 NLSLGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNR 180

Qy    292 ASFSQYGAGLDIVAPGVNVQSTYPGSTYASFNGTSMATPHVGVAALVKQKNPSWSNVQI 351
              |||||||||||||||||||||||||||||:|||||||||| :|||||||||||||||||
Db    181 ASFSQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKNPSWSNVQI 240

Qy    352 RNHLKNTATSLGSTNLYGSGLVNAEAATR 380
              |||||||:|||||||||||||||||||||
Db    241 RNHLKNTATSLGSTNLYGSGLVNAEAATR 269
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2006, 18:25:57 ; Search time 48 Seconds
(without alignments)
654.516 Million cell updates/sec

Title: US-09-931-701A-2
Perfect score: 1901
Sequence: 1 MNKPLGKIVASTALLISVAF.........SLGSTNLYGSGLVNAEAATR 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
Issued_Patents_AA:*
1: /cgn2_6/ptcdata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptcdata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptcdata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptcdata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptcdata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptcdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1777	93.5	380	2	US-09-049-867-2	Sequence 2, Appli
2	1777	93.5	380	2	US-08-269-050-2	Sequence 2, Appli
3	1777	93.5	380	6	5217878	Patent No. 5217878
4	1777	93.5	380	6	5336611-2	Patent No. 5336611-2
5	1772	93.2	380	1	US-07-918-318-2	Sequence 2, Appli
6	1772	93.2	380	2	US-08-411-724-2	Sequence 2, Appli
7	1772	93.2	380	1	US-08-853-494-2	Sequence 2, Appli
8	1772	93.2	380	2	US-08-716-293-1	Sequence 1, Appli
9	1772	93.2	380	2	US-09-445-270-7	Sequence 7, Appli
10	1770	93.1	380	2	US-09-824-893A-261	Sequence 261, App
11	1758	92.5	380	2	US-09-512-251A-4	Sequence 4, Appli
12	1618	85.1	349	1	US-07-661-378A-2	Sequence 2, Appli
13	1491	78.4	378	1	US-07-772-087-3	Sequence 3, Appli
14	1284	67.5	269	1	US-08-431-387-6	Sequence 6, Appli
15	1284	67.5	269	4	US-08-269-050-4	Sequence 4, Appli
16	1284	67.5	269	1	US-09-090-207-1	Sequence 1, Appli
17	1284	67.5	269	2	US-09-512-251A-4	Sequence 4, Appli
18	1284	67.5	269	2	US-09-515-150A-4	Sequence 4, Appli
19	1284	67.5	269	2	US-09-196-281-7	Sequence 7, Appli
20	1284	67.5	269	2	US-10-336-324-4	Sequence 4, Appli
21	1284	67.5	269	6	5217878-3	Patent No. 5217878
22	1282	67.4	269	1	US-07-706-691G-16	Sequence 16, Appli
23	1282	67.4	269	1	US-08-254-021-16	Sequence 16, Appli
24	1282	67.4	269	1	US-08-618-446-16	Sequence 16, Appli
25	1282	67.4	269	1	US-08-980-135-16	Sequence 16, Appli
26	1282	67.4	269	2	US-09-585-798-16	Sequence 16, Appli
27	1279	67.3	269	1	US-07-706-691G-31	Sequence 31, Appli
28	1279	67.3	269	1	US-08-254-021-31	Sequence 31, Appl
29	1279	67.3	269	1	US-08-431-387-5	Sequence 5, Appl
30	1279	67.3	269	1	US-08-322-677A-10	Sequence 10, Appl
31	1279	67.3	269	1	US-08-322-676-10	Sequence 10, Appl
32	1279	67.3	269	1	US-08-140-083A-10	Sequence 10, Appl
33	1279	67.3	269	1	US-08-618-446-31	Sequence 31, Appl
34	1279	67.3	269	2	US-08-898-218-10	Sequence 10, Appl
35	1279	67.3	269	2	US-08-848-793-10	Sequence 31, Appl
36	1279	67.3	269	2	US-09-255-502-5	Sequence 31, Appl
37	1279	67.3	269	2	US-08-090-207-2	Sequence 5, Appl
38	1279	67.3	269	3	US-09-024-532-3	Sequence 3, Appl
39	1279	67.3	269	2	US-09-178-155-6	Sequence 5, Appl
40	1279	67.3	269	2	US-09-445-270-2	Sequence 2, Appli
41	1279	67.3	269	2	US-09-467-536A-1	Sequence 1, Appli
42	1279	67.3	269	2	US-09-234-957-1	Sequence 1, Appli
43	1279	67.3	269	2	US-09-585-798-31	Sequence 31, Appl
44	1279	67.3	269	2	US-09-417-359A-3	Sequence 3, Appl
45	1279	67.3	269	1	US-09-049-867-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-049-867-2
Sequence 2, Application US/09049867
Patent No. 6124097
GENERAL INFORMATION:
APPLICANT: Van Eekelen, Christiaan A.G.
APPLICANT: Van Der Laan, Johannes C.
APPLICANT: Mulleners, Leo J.S.M.
TITLE OF INVENTION: Stable Gene Amplification in
TITLE OF INVENTION: Chromosomal DNA of Prokaryotic Microorganisms
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/049,867
FILING DATE: 27-MAR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/295,082
FILING DATE: 24-AUG-94
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-049-867-2

Query Match 93.5%; Score 1777; DB 2; Length 380;
Best Local Similarity 93.2%; Pred. No. 4.1e-144;
Matches 354; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

```
Qy   1 MNKPLGKIVASTALLISVAFSSSTASAAEBAKEKYLIGFNEQEAVSEFVEQVDANNDVAV  60
        | ||||||||||||||||||||||||||||||||||||||||||||||||||| ||:
Db   1 MKKPLGKIVASTALLISVAFSSSTASAAEBAKEKYLIGFNEQEAVSEFVEQVEANDEVAI  60

Qy  61 LSEBEEVEIELLHHFETIPVLSVELSPEDVDTLELDPAISYIEEDVEVSIMAQSVPWGIS 120
        ||||||||||||||||||||||||||||||||| |||:|||||||||||:||||||||||
Db  61 LSEBEEVEIELLHHFETIPVLSVELSPEDVDALELDPAISYIEEDABVTTMAQSVPWGIS 120
```

OM protein - protein search, using sw model

GenCore version 5.1.7
Copyright (c) 1993 - 2006   Bioceleration Ltd.

Run on: March 13, 2006, 18:18:46 ; Search time 224 Seconds
(without alignments)
1196.878 Million cell updates/sec

Title:
Perfect score: 1901
Sequence: 1 MNKPLGKIVASTALLISVAF.........SLGSTNLYGSGLVNAEAATR 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Accession	Description
1	1777	93.5	380	1	ELYA_BACAO	P27693	bacillus al
2	1772	93.2	380	1	ELYA_BACCS	P41362	bacillus cl
3	1768	93.0	380	1	PRTM_BACSK	Q99405	bacillus cl
4	1500.5	78.9	379	2	O66153_BACSP	O66153	bacillus sp
5	1494	78.6	379	2	ELYA_BACYA	P27724	bacillus sp
6	1387	73.0	382	2	Q45522_9BACI	Q45522	bacillus sp
7	1279	67.3	269	1	SUBS_BACLE	P29600	bacillus le
8	1278	67.2	269	1	SUBS_BACLE	P29599	bacillus le
9	1041.5	54.8	361	2	Q45521_BACSP	Q45521	bacillus sp
10	1037.5	54.6	361	1	ELYA_BACHD	P41363	bacillus ha
11	978.5	51.5	374	1	Q45523_BACSP	Q45523	bacillus sp
12	976.5	51.4	378	2	O45466_BACSP	O45466	bacillus sp
13	974	51.2	376	2	Q76KL9_9BACI	Q76KL9	bacillus sp
14	949.5	49.9	381	2	Q58GF0_BACSU	Q58gf0	bacillus su
15	945	49.7	381	2	Q847A2_BACSU	Q847a2	bacillus su
16	942.5	49.6	381	1	SUBN_BACNA	P35835	bacillus su
17	942.5	49.6	381	1	SUBN_BACNA	Q54EF3	bacillus su
18	940.5	49.5	381	1	Q84F18_BACSU	Q84f18	bacillus su
19	937.5	49.3	381	1	SUBT_BACSA	P00783	bacillus su
20	937.5	49.3	381	1	O58LV1_BACSU	O58lv1	bacillus su
21	934.5	49.2	381	1	SUBT_BACSU	P04189	bacillus su
22	933	49.1	381	1	SUBT_BACST	P29142	bacillus st
23	933.5	49.1	381	1	O58LU9_BACSU	O58lu9	bacillus su
24	932	49.0	382	1	SUBT_BACAM	P00782	bacillus am
25	930	48.9	382	2	Q5XZF9_9BACL	Q5xzf9	brevibacill
26	928.5	48.8	382	2	Q5EFD9_BACNA	Q5efd9	bacillus su
27	926	48.7	382	2	Q6IT79_9BACI	Q6it79	bacillus sp
28	920.5	48.4	381	2	Q7WVA6_BACSU	Q7wva6	bacillus su
29	918.5	48.3	377	2	Q6L9T7_9BACI	Q6l9t7	bacillus sp
30	918.5	48.3	377	2	Q9FDF4_BACLI	Q9fdf4	bacillus li
31	918.5	48.3	379	2	Q65LP7_BACLD	Q65lp7	bacillus li
32	913	48.0	382	2	O87655_BACSU	O87655	bacillus su
33	910	47.9	382	2	Q5UKQ4_9BACI	Q5ukq4	bacillus sp
34	909.5	47.8	379	1	SUBT_BACLI	P00780	bacillus li
35	909.5	47.8	379	2	Q45467_BACSP	Q45467	bacillus sp
36	898.5	47.3	379	2	Q6PNN5_BACLI	Q6pnn5	bacillus li
37	894.5	47.1	379	2	Q53521_BACLI	Q53521	bacillus li
38	894.5	47.1	379	2	Q6BCN9_BACMO	Q6bcn9	bacillus mo
39	894.5	47.1	379	2	Q4PKR6_BACLI	Q4pkr6	bacillus li
40	885.5	46.6	374	2	Q9F942_BACLI	Q9f942	bacillus li
41	885.5	46.6	374	2	Q9F941_BACLI	Q9f941	bacillus li
42	883	46.4	404	2	Q76L29_9BACI	Q76l29	bacillus sp
43	881.5	46.2	374	2	Q9F943_BACLI	Q9f943	bacillus li
44	878	46.2	379	2	Q45301_BACLI	Q45301	bacillus li
45	867.5	45.6	310	2	Q9FDF3_BACLI	Q9fdf3	bacillus li

ALIGNMENTS

RESULT 1
ELYA_BACAO

```
ID   ELYA_BACAO     STANDARD;      PRT;   380 AA.
AC   P27693;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Alkaline protease precursor (EC 3.4.21.-).
OS   Bacillus alcalophilus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1445;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=91282483; PubMed=2059048;
RA   van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA   Quax W.J.;
RT   "Cloning, characterization, and multiple chromosomal integration of a
RT   Bacillus alkaline protease gene.";
RL   Appl. Environ. Microbiol. 57:901-909(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC   STRAIN=PB92;
RX   MEDLINE=92390330; PubMed=1518788;
RA   van der Laan J.C., Teplyakov A.V., Kelders H., Kalk K.H., Misset O.,
RA   Mulleners L.J.M., Dijkstra B.W.;
RT   "Crystal structure of the high-alkaline serine protease PB92 from
RT   Bacillus alcalophilus.";
RL   Protein Eng. 5:405-411(1992).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RC   STRAIN=PB92;
RX   MEDLINE=93078250; PubMed=1447775;
RA   Sobek H., Hecht H.-J., Aehle W., Schomburg D.;
RT   "X-ray structure determination and comparison of two crystal forms of
RT   a variant (Asn115Arg) of the alkaline protease from Bacillus
RT   alcalophilus refined at 1.85-A resolution.";
RL   J. Mol. Biol. 228:108-117(1992).
RN   [4]
RP   STRUCTURE BY NMR OF 112-380.
RC   STRAIN=PB92;
RX   MEDLINE=97277237; PubMed=9115441; DOI=10.1016/S0969-2126(97)00208-6;
RA   Martin J.R., Mulder F.A., Karimi-Nejad Y., van der Zwan J.,
RA   Mariani M., Schipper D., Boelens R.;
RT   "The solution structure of serine protease PB92 from Bacillus
RT   alcalophilus presents a rigid fold with a flexible substrate-binding
RT   site.";
RL   Structure 5:521-532(1997).
CC   -!- COFACTOR: Binds 2 calcium ions per subunit.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the peptidase S8 family.
CC   -----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
                    GenCore version 5.1.7
            Copyright (c) 1993 - 2006  Biocceleration Ltd.


OM protein - protein search, using sw model

Run on:          March 13, 2006, 18:18:21 ; Search time 189 Seconds
                                            (without alignments)
                                            883.407 Million cell updates/sec

Title:           US-09-931-701A-2
Perfect score:   1901
Sequence:        1 MNKPLGKIVASTALLISVAF.........SLGSTNLYGSGLVNAEAATR 380

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched:        2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters:   2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_21:*
                 1: geneseqp1980s:*
                 2: geneseqp1990s:*
                 3: geneseqp2000e:*
                 4: geneseqp2001e:*
                 5: geneseqp2002e:*
                 6: geneseqp2003aa:*
                 7: geneseqp2003ba:*
                 8: geneseqp2004s:*
                 9: geneseqp2005e:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1901	100.0	380	5	AAU78181	Aau78181 B. clausi
2	1777	93.5	380	1	AAP80867	Aap80867 Sequence
3	1777	93.5	380	1	AAP90915	Aap90915 PB92 seri
4	1777	93.5	380	4	AAR43046	Aar43046 Wild-type
5	1777	93.5	380	4	AAB82790	Aab82790 Bacillus
6	1777	93.5	380	5	AAE16857	Aae16857 Bacillus
7	1777	93.5	380	7	ADE81100	Ade81100 Bacillus
8	1772	93.2	380	2	AAR10892	Aar10892 B.alcalop
9	1772	93.2	380	2	AAR27131	Aar27131 Stabilise
10	1772	93.2	380	2	AAR56514	Aar56514 High alka
11	1772	93.2	380	2	AAR77377	Aar77377 Bacillus
12	1772	93.2	380	2	AAW82797	Aaw82797 B. lentus
13	1772	93.2	380	3	AAY77004	Aay77004 Bacillus
14	1772	93.2	380	3	AAB03776	Aab03776 Subtilisi
15	1772	93.2	380	7	ADE81101	Ade81101 Bacillus
16	1772	93.2	380	8	ADJ46880	Adj46880 B. lentus
17	1770	93.1	380	4	AAU38874	Aau38874 Subtilisi
18	1769	93.1	380	2	AAR31928	Aar31928 High alka
19	1768	93.0	380	2	AAR64214	Aar64214 Alkaline
20	1768	93.0	380	2	AAR94075	Aar94075 Bacillus
21	1768	93.0	380	2	AAR94080	Aar94080 Bacillus
22	1767	93.0	380	2	AAR34611	Aar34611 High alka
23	1767	93.0	380	2	AAR94081	Aar94081 Bacillus
24	1766	92.9	380	2	AAR11189	Aar11189 Pre-pro a
25	1766	92.9	380	2	AAR34610	Aar34610 High alka
26	1766	92.9	380	2	AAR75163	Aar75163 B. lentus
27	1766	92.9	380	2	AAR94076	Aar94076 Bacillus
28	1766	92.9	380	2	AAR94077	Aar94077 Bacillus
29	1766	92.9	380	2	AAR94083	Aar94083 Bacillus
30	1766	92.9	380	2	AAR94078	Aar94078 Bacillus
31	1766	92.9	380	2	AAR94079	Aar94079 Bacillus
32	1766	92.9	380	2	AAW34891	Aaw34891 Preferred
33	1766	92.9	380	3	AAW34774	Aaw34774 Preferred
34	1766	92.9	380	2	AAY77005	Aay77005 Bacillus
35	1766	92.9	380	2	AAB03777	Aab03777 Subtilisi
36	1766	92.8	380	2	ADJ46881	Adj46881 B. lentus
37	1765	92.8	380	2	AAR34615	Aar34615 High alka
38	1765	92.8	380	2	AAR34605	Aar34605 High alka
39	1764	92.8	380	2	AAR34612	Aar34612 High alka
40	1764	92.8	380	2	AAR94082	Aar94082 Bacillus
41	1763	92.7	380	2	AAR34618	Aar34618 High alka
42	1763	92.7	380	2	AAR34613	Aar34613 High alka
43	1763	92.7	380	2	AAR34614	Aar34614 High alka
44	1763	92.7	380	2	AAR34617	Aar34617 High alka
45	1763	92.7	380	2	AAR34620	Aar34620 High alka

ALIGNMENTS

```
RESULT 1
AAU78181
AAU78181;

ID   AAU78181 standard; protein; 380 AA.
XX
AC   AAU78181;
XX
DT   29-AUG-2003  (revised)
DT   05-JUN-2002  (first entry)
XX
DE   B. clausii subtilase.
XX
KW   Subtilase; detergent; laundry; dishwash; egg stain removal.
XX
OS   Bacillus clausii; strain HSB10.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..111
FT                   /label= Signal_peptide
FT   Protein         112..380
FT                   /label= Mature_subtilase
XX
DR   WPI; 2002-280919/32.
DR   N-PSDB; ABK12132.
XX
PN   WO200216547-A2.
XX
PD   28-FEB-2002.
XX
PF   21-AUG-2001; 2001WO-DK000551.
XX
PR   21-AUG-2000; 2000DK-00001232.
XX
PA   (NOVO ) NOVOZYMES AS.
XX
PI   Outrup H, Pedersen PE, Sorensen MV;
XX
PT   Novel subtilase enzyme that exhibits low inhibition by substances present
PT   in eggs such as trypsin inhibitor type IV-O, useful in a cleaning or
PT   detergent composition and for removal of egg stains.
XX
PS   Claim 3; Page 87-88; 88pp; English.
XX
CC   The invention relates a subtilase enzyme and its encoding nucleic acid
CC   (or a protein 95 % similar to the subtilase). Also included are a
CC   construct comprising the nucleic acid operably linked to one or more
CC   control sequences capable of directing the expression of the subtilase in
CC   a suitable host, a recombinant expression vector comprising the construct
```

OM protein - protein search, using sw model

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: March 13, 2006, 18:22:07 ; Search time 41 Seconds
891.765 Million cell updates/sec
(without alignments)

Title: US-09-931-701A-2
Perfect score: 1901
Sequence: 1 MNKPLGKIVASTALLISVAF...........SLGSTNLYGSGLVNAEAATR 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1777	93.5	380	2	A49778	high-alkaline seri
2	1494	78.6	378	2	A33973	high-alkaline seri
3	1387	73.0	382	2	I39780	subtilisin (EC 3.4
4	1044.5	54.9	361	2	A48373	high-alkaline seri
5	1037.5	54.6	361	2	G83756	subtilisin-type al
6	978.5	51.5	374	2	I39781	subtilisin (EC 3.4
7	942.5	49.6	381	2	JH0778	subtilisin (EC 3.4
8	938.5	49.4	381	2	SUBSI	subtilisin (EC 3.4
9	937.5	49.3	381	2	SUBSS	subtilisin (EC 3.4
10	933.5	49.1	381	2	JQ1487	subtilisin (EC 3.4
11	932	49.0	382	1	SUBSN	subtilisin (EC 3.4
12	909.5	47.8	379	1	SUBSCL	subtilisin (EC 3.4
13	843	44.3	272	2	A23624	subtilisin (EC 3.4
14	821.5	43.2	372	2	D83735	subtilisin-type al
15	821.5	43.2	275	2	JC1085	subtilisin (EC 3.4
16	796.5	41.9	274	1	SUBSD	subtilisin (EC 3.4
17	681.5	35.8	645	1	SUBSMP	serine proteinase
18	642.5	33.8	397	2	JW0075	cysteine-dependent
19	618	32.5	384	2	JC4802	alkaline proteinas
20	603	31.7	279	1	SUMYTV	thermitase [EC 3.4
21	559.5	29.4	436	1	SUMYTV	alkaline serine pr
22	553.5	29.1	401	2	I39974	serine proteinase
23	553.5	29.1	525	2	G84406	halolysin [importe
24	539	28.4	792	2	H83736	subtilisin-type al
25	534	28.1	440	2	H72784	probable alkaline
26	528.5	27.8	488	2	AI1930	proteinase [import
27	519.5	27.3	326	1	S75976	hypothetical prote
28	514.5	27.1	326	1	C41335	microbial serine p
29	512	26.9	580	2	S11890	serine proteinase

ALIGNMENTS

RESULT 1
A49778
high-alkaline serine proteinase (EC 3.4.21.-) precursor - Bacillus alcalophilus
A;Alternate names: subtilisin homolog, high-alkaline
C;Species: Bacillus alcalophilus
C;Date: 12-Mar-1994 #sequence_revision 24-Feb-1995 #text_change 05-Oct-2004
C;Accession: A49778; JC1244
R;van der Laan, J.C.; Gerritse, G.; Mulleners, L.J.S.M.; van der Hoek, R.A.C.; Quax, W
Appl. Environ. Microbiol. 57, 901-909, 1991
A;Title: Cloning, characterization, and multiple chromosomal integration of a Bacillus
A;Experimental source: strain PB92, ATCC 31408
A;Reference number: A49778; MUID:9128249; PMID:2059048
A;Accession: A49778
A;Molecule type: DNA
A;Residues: 1-380 <VAN>
A;Cross-references: UNIPROT:P27693; UNIPARC:UPI000002BADA; GB:M65086; NID:g142456; PID
A;Experimental source: strain PB92, ATCC 31408
A;Reference number: A49778; MUID:9128249; PMID:2059048
R;Takami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Aono, R.; Hori
Biosci. Biotechnol. Biochem. 56, 1455-1460, 1992
A;Title: Molecular cloning, nucleotide sequence, and expression of the structural gene
A;Reference number: JC1244; MUID:93043753; PMID:1368952
A;Molecule type: DNA
A;Residues: 1-195; 'S', 197-380 <TAK>
A;Cross-references: UNIPARC:UPI0000002CPSA; GB:D13157; NID:g2162231; PIDN:BAA02442.1; PI
A;Experimental source: UNIPARC:UPI0000002CPSA; GB:D13157; NID:g2162231; PIDN:BAA02442.1; PI
A;Experimental source: Bacillus alcalophilus Vedder, ATCC 21522 (Bacillus sp. 221)
C;Superfamily: Subtilisin; subtilisin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-111/Domain: activation peptide #status predicted <PRO>
F;112-380/Product: alkaline serine proteinase #status predicted <MAT>
F;134-340/Domain: subtilisin homology <SBT>
F;143,173,326/Active site: Asp, His, Ser #status predicted

Query Match 93.5%; Score 1777; DB 2; Length 380;
Best Local Similarity 93.2%;
Matches 354; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

Qy   1 MNKPLGKIVASTALLISVAFSSSISASAAEEAKEKYLIGFNEQEAVSEFVEQVDANNDVAV 60
         ||||||||||||||||||||||||||| |||| |||||||||||||||||||| | |
Db   1 MKKPLGKIVASTALLISVAFSSSISASAAEEAKEKYLIGFNEQEAVSEFVEQVEANDEVAI 60

Qy   61 LSEEEVEIELLLHEFETIPVLSVELSPEDVDTLELDPAISYIEEDVEVSIMAQSVPWGIS 120
         |||||| || ||||||||||||||| |||||||||||||||||||||||||||||||||
Db   61 LSEEEVEIELLHEFETIPVLSVELSPEDVDALELDPAISYIEEDAEVTTMAQSVPWGIS 120

Qy 121 RVQAPAAHNRGVTGSGVKVAVLDTGISAHPDLNIRGGASFVTGEPTYQDGNGHGTHVAGT 180
         ||||||||||| ||||||||||||||| |||||| |||||| | | |||||||||||||
Db 121 RVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQDGNGHGTHVAGT 180

Qy 181 IAALNNSIGVLGVAGPNAELYAVKVLAANGRGPVSSIAQGLEWAGNNGMDVANLSLGSPSP 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 [continues] 240